Kevin T. Morgan, GlaxoWellcome, Inc.

Research Proposal for the Toxicology and Risk Assessment Breakout Group

Keep the 'Bio' in 'Bioinformatics!

As a consequence of the 'omics revolution,' developing the necessary balance of mathematical and 'intuitive biology' skills in toxicologists and risk assessors of the future will need to be planned and adequately resourced.

Proposal:

Invest in training programs for toxicologists working with data sets derived from genomics, transcriptomics, proteomics, physiomics, metabolomics, etc.

The statement that "new pieces of technology commonly give rise to information overload..." (Nicholls 1999) is clearly true for the revolution taking place in molecular biology technology. Complete sequencing of human and other genomes, availability of large-scale gene expression arrays, with an ever increasing number of genes displayed, and the steady improvement in large-scale protein expression technology, are starting to change our approaches to biology, including the field of toxicology (Waring and Ulrich 2000). It evident that the rate of generation of data points exceeds our ability to confirm them at anything but a trivial level, while interpretation presents an even greater challenge. A relatively small gene expression data set, such as that recently published on the response of pancreatic β-cells to high glucose levels (Webb, Akbar et al. 2000), sends one searching through texts on intermediary metabolism when attempting to put these data in perspective. Considerable 'on line' assistance may be obtained with interpretation, if on can negotiate links to a dizzying array of databases, such KEGG (http://www.genome.ad.jp/kegg/metabolism.html) or EPD (http://www.epd.isb-sib.ch/). You, personally, still need to understand these pathways and regulatory circuits before a

PM3006518346

Source: https://www.industrydocuments.ucsf.edu/docs/rygj0001

meaningful interpretation is possible, whether your gene data sets are well annotated or not.

Interpretation of these large data sets falls within the reach of the rapidly evolving science/art of bioinformatics. The complexity and scope of these data sets tends to overwhelm all but the most dedicated observer. As a direct consequence of the innate power of mathematics, we (the biologists) run the risk of allowing mathematical approaches (Baldi and Søren 1998) to mold and define the discipline of bioinformatics. For many years (it appeared to me) statisticians dominated the field of Human Disease Risk Assessment, while the data they used were largely derived from applied biology. Dominance can be an unhealthy thing (though I have a growing respect for statistics), while a healthy balance of power can be difficult to achieve. If the biologists hope to exert appropriate levels of influence on the application of 'omics' to toxicology and human disease risk assessment, they must find ways to train a new generation of toxicologists and risk assessors. This new generation will have the interesting challenge of merging the massive data sets derived from new technologies with tried and true disciplines, such as anatomy, physiology, biochemistry, pharmacology, and pathology. The 'old' disciplines are easy to forget when you are carried away by the (largely) as yet unfulfilled promise of these new technologies.

References

Baldi, P. and B. Søren (1998). Bioinformatics, The Machine Learning Approach. Cambridge, Massachusetts, The MIT Press.

Nicholls, P. (1999). The mitochondrial and bacterial respiratory chains: From MacMunn and Keilin to current concepts. Frontiers of Cellular Bioenergetics. F. G. a. J. M. T. Sergio Papa. New York, Kluwer Academic/Plenum Publishers: 1-22.

Waring, J. F. and R. G. Ulrich (2000). "The impact of genomics-based technologies on drug safety evaluation." Annual Reviews in Pharmacology and Toxicology **40**: 335-352.

Webb, G. C., M. S. Akbar, et al. (2000). "Expression profiling of pancreatic β cells: glucose regulation of secretory and metabolic pathway genes." Proceedings of the National Academy of Sciences, USA **97**(11): 5773-5778.

Source: https://www.industrydocuments.ucsf.edu/docs/rygj0001